



Predicting Microevolutionary Responses to Directional Selection on Heritable Variation

Peter R. Grant; B. Rosemary Grant

Evolution, Vol. 49, No. 2. (Apr., 1995), pp. 241-251.

Stable URL:

<http://links.jstor.org/sici?sici=0014-3820%28199504%2949%3A2%3C241%3APMRTDS%3E2.0.CO%3B2-X>

Evolution is currently published by Society for the Study of Evolution.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ssevol.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

PREDICTING MICROEVOLUTIONARY RESPONSES TO DIRECTIONAL SELECTION ON HERITABLE VARIATION

PETER R. GRANT AND B. ROSEMARY GRANT

Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544-1003

E-mail: PRGrant@PUCC.Princeton.edu

Abstract.—Microevolution of quantitative traits in the wild can be predicted from a knowledge of selection and genetic parameters. Testing the predictions requires measurement of the offspring of the selected group, a requirement that is difficult to meet. We present the results of a study of Darwin's finches on the Galápagos island of Daphne Major where this requirement is met. The study demonstrates microevolutionary consequences of natural selection.

The population of medium ground finches, *Geospiza fortis*, experienced size-selective mortality during a drought in 1976–1977; large birds with deep beaks survived better than small birds. During another drought, 1984–1986, the population experienced selection in the opposite direction on beak traits. Changes in food supply were the apparent causes of selection on beak traits in both episodes. As expected from the high heritabilities of all measured traits, the effects of selection were transmitted to the next generation. Evolutionary responses to both episodes of selection were quantitatively well predicted in general. This allows us to conclude that, to a first approximation, targets of selection were identified correctly, and genetic parameters were correctly estimated. Nevertheless, not all responses of individual traits were equally well predicted. A search for possible reasons for the largest discrepancies revealed evidence of bias caused by environmental effects on growth and adult size of some traits, as well as possible selection on the offspring generation before their measurement.

These findings illustrate an important assumption in the study of microevolution: that the environments experienced during growth to maturity by the parental and offspring generations are the same, for otherwise a measured difference between generations may have a partly environmental cause, thereby giving a misleading estimate of the evolutionary response to selection. Simple extrapolations from observed selection to long-term evolution may underestimate the total force of selection involved if it oscillates in direction or acts on the environmental variance.

Key words.—Beaks, body size, Darwin's finches, extrapolation, Galápagos, genetic correlations, multivariate evolution, nutrition.

Received December 15, 1993. Accepted May 23, 1994.

No scientific theory is worth anything unless it enables us to predict something which is actually going on. Until that is done, theories are a mere game of words, and not such a good game as poetry.

—J. B. S. Haldane

It is often said that scientific theories are useful only to the extent that they are predictive. The best scope for a predictive theory of evolution lies in the area of genetics, because the mathematical machinery has been developed for the precise prediction of evolutionary change caused by selection.

The standard equation for predicting an evolutionary response to selection on quantitative traits was developed for the purpose of improving animal and plant yields under artificial selection (Bulmer 1980; Falconer 1989). The response (R) in a trait is *predicted* by the product of the strength of selection (s) on that trait and its narrow-sense heritability (h^2). The response is *measured* as the difference between the mean of the trait in the group prior to selection and the mean of the trait in the next generation. It is typically expressed in units of standard deviations of the initial group.

Natural selection is likely to affect more than one trait simultaneously, and traits are correlated to some degree; therefore, a multivariate predictive equation is needed for the study of natural selection in the wild. An equation for this purpose was developed by Lande (1979). It predicts an evolutionary response to selection on a trait in terms of not only the strength of selection on the trait and its genetic variance, but also selection intensities on correlated traits and the corresponding genetic covariances.

The multivariate equation has a usefulness in two domains:

in understanding how natural selection may have worked in the past in the transformation of morphological relations when one species evolved into another; and in the study of microevolution as a contemporary, observable, and measurable process. Lande (1979) illustrated its usefulness in the first context by applying it to the problem of how mammalian brain size changed in relation to body size during the course of species-, genus-, and ordinal-level evolution. It has been adopted as a tool in other contexts, for example, by Reymont (1983) in the study of Cretaceous microfossils, by Price et al. (1984a) in the study of speciation, and by Schluter (1984) in the reconstruction of phylogeny. As a tool for the study of microevolution, it has a rich potential that has been exploited mainly in theoretical studies, for example, in examining the evolution of genetically correlated growth characteristics that affect sets of functionally related adult traits (Cheverud 1984; Kingsolver and Wiernasz 1991; Kirkpatrick and Lofsvold 1992). In this paper we illustrate its empirical potential with the results of a study of microevolution in Darwin's finches on the Galápagos island of Daphne Major.

We use the term *natural selection* to mean a systematic variation in fitness (survival or reproduction) among individuals in a population associated with variation in a trait or traits. To the degree that the trait varies in a heritable manner, the effects of selection will be passed on to the next generation as an evolutionary response. Thus, in our language,

natural selection occurs within a generation and evolution occurs between generations.

It would appear to be a straightforward matter to measure a group of individuals before and again after selection, and then measure the offspring of the selected sample at the same, usually adult, stage of development, before any selection on them. This would yield a measurement of the evolutionary response to selection, and it could be compared with a prediction of the response based on prior estimation of the appropriate genetic parameters. In practice, it is often difficult to identify the offspring of particular individuals of the selected group, capture them, and measure them. Paternity assessment gives rise to serious problems, especially for the study of plants, and even when that is not a problem finding and capturing a sufficient sample of the offspring is a problem, especially for the study of animals that often disperse over long distances or live in cryptic (soil) or otherwise largely inaccessible (marine) habitats. For these and other reasons, there is a dearth of studies of evolutionary responses to selection in nature (Flux and Flux 1982), even though selection itself has been measured on numerous occasions (Endler 1986; Price and Boag 1987; and many references thereafter).

These problems are minimized in the study of bird populations on small islands, where most if not all individuals can be uniquely marked and their reproductive fates followed (Grant 1986; Schluter and Smith 1986a,b; van Noordwijk and van Balen 1988; Gustafsson 1989). The study of Darwin's finches on Daphne Major island is an exemplar. Over an 18-mo period of drought and no breeding, from the middle of 1976 to the beginning of 1978, *G. fortis* was subjected to directional selection; large birds survived better than small birds (Boag and Grant 1981, 1984a). Selection-gradient analysis (Lande and Arnold 1983) identified body size and beak depth as the targets of selection (Price et al. 1984b). From mid-1984 to the beginning of 1986, the population was again subjected to selection in association with heavy mortality (and no breeding) during a drought, but this time in the opposite direction, and beak traits were the only targets of selection (Gibbs and Grant 1987a; Grant and Grant 1993). Changes in food supply were the apparent causes of selection on beak traits in both episodes (Boag and Grant 1981, 1984a; Gibbs and Grant 1987a; Grant and Grant 1993). Different changes produced selection pressures in different directions.

Evolutionary responses to selection are to be expected because heritabilities of all measured traits are high (Boag and Grant 1978; Boag 1983; Grant and Grant 1994). Boag (1983) showed that offspring of the survivors of the first selection episode, born in 1978, were larger on average than the group before selection in 1976, but did not attempt a multivariate analysis of the response. The only multivariate analysis performed so far was made with data from the second selection episode. The response observed in 1987 to selection in 1984–1986 corresponded closely to that predicted from Lande's (1979) equation for multivariate evolution of quantitative traits (Grant and Grant 1993). However, the analysis was restricted to the two beak traits identified as statistically significant targets of selection and, as such, was less than complete and potentially flawed.

Here we present the results of an analysis that predicts the evolutionary responses to selection in both episodes using

the complete set of data of six measured morphological traits and survival as the sole measure of fitness. We show that both responses are well predicted. This allows us to conclude that, to a first approximation, targets of selection have been correctly identified and genetic parameters have been correctly estimated. If any targets were inadvertently omitted from the analysis, they must have been highly correlated with the identified targets (see Price and Langen 1992). We then consider why the responses of some traits were less well predicted than others. A search for possible causes reveals evidence of bias arising from effects of the environment preventing attainment of full adult size and selection on the offspring generation before their measurement. Another possible source of bias, sexual selection on the survivors at the time of breeding, is found to be not applicable. These findings illustrate an important assumption in the study of microevolution: that the environments experienced during growth to maturity by the parental and offspring generations are the same, for otherwise a measured difference between generations may have an entirely environmental cause and falsely indicate evolution.

METHODS

The island, finches, and study methods have been described in detail in earlier publications (e.g., Abbott et al. 1977; Boag and Grant 1984a,b; Grant 1986). Daphne Major is a small island (0.34 km²) in the center of the Galápagos archipelago. It supports breeding populations of four species of Darwin's finches: *Geospiza fortis* (medium ground finch, ≈17 g), *G. fuliginosa* (small ground finch, ≈14 g), *G. magnirostris* (large ground finch, ≈30 g), and *G. scandens* (cactus finch, ≈21 g). Finches have been studied every year from 1973 to 1993. Over the period 1976–1992, harmonic mean breeding-population sizes were 198 *G. fortis*, 80 *G. scandens*, 3 *G. fuliginosa*, and, from 1983 onward, 6 *G. magnirostris* (Grant 1993; Grant and Grant 1995).

Adult and immature finches were trapped in mist nets each year, usually in January and February at the end of the dry, non-breeding, season. They were measured, banded and released. Birds were weighed to the nearest 0.1 g with a Pesola® spring balance; wing length was measured with a wing ruler to the nearest millimeter; tarsus length and bill length were measured with dividers, and bill depth and bill width were measured with callipers, all to the nearest 0.1 mm. Most birds were measured by P.R.G., and correction factors, calculated from birds measured by P.R.G. and other measurers, were applied to the rest. Repeated measurements of the same bird were averaged. Repeatabilities are high, especially for bill measurements (Grant and Grant 1994). Birds were given a unique combination of four bands, two on each leg: one numbered metal band and three colored plastic (PVC) bands coded to correspond to the number on the metal band. Nestlings were banded at an age of 8 d, and measured upon capture in mist nets when fully grown (60 d or older; Boag 1984). Parents were identified by observation of adults feeding nestlings.

Analyses

Selection analyses were performed on ln-transformed data, standardized by dividing by the standard deviation (Lande

and Arnold 1983; Price et al. 1984b; Gibbs and Grant 1987a). Before analysis, weight was converted to the cube root of weight. The net effect of selection on a trait, the selection differential (s), was calculated as the difference in the mean value of the trait before and after selection. The direct effect of selection on each trait, or selection gradient (β), which is independent of selection on correlated traits, was estimated by the partial-regression coefficient of relative fitness (survival) on each trait. Statistical significance of selection coefficients was assessed by t -tests. Significance of the gradients was confirmed by t -tests of the coefficients from a logistic regression analysis (Grant and Grant 1993). An evolutionary response to selection was calculated on the untransformed scale as the difference between the mean of a trait before selection and the mean of the same trait in the offspring of the selected group (Falconer 1989).

Observed evolutionary responses were compared with those predicted from estimates of standardized selection coefficients, heritabilities, and genetic correlations (Lande 1979). The expected shift in each trait mean (\bar{z}_i) was calculated from:

$$\Delta\bar{z}_i = \beta_1 \cdot h_1^2 + \beta_2 \cdot h_2 \cdot h_1 \cdot r_{a21} \dots \beta_6 \cdot h_6 \cdot h_1 \cdot r_{a61}.$$

β_{1-6} are entries in the selection gradient of the direct effects of selection on each of the traits 1–6, h^2 is the heritability of a trait, h is its square root, and r_a is the genetic correlation between two traits. Thus, for each trait, the predicted response is the product of direct selection on that trait and indirect selection on all of the other five correlated traits. Heritabilities and genetic correlations were calculated from 1976 data supplemented by a few extra families formed in 1978 (Boag 1983) and a few extra measurements made since then. On the basis of the magnitude of the standard errors, the 1976 sample alone (Boag and Grant 1978) was considered too small for the estimation of these genetic parameters.

Sampling errors for the predictions were obtained by bootstrapping, performed by D. Schluter (pers. comm. 1994). The 41 families of parents and offspring were resampled randomly (with replacement) to yield 500 replicates of the genetic covariance matrix. A β vector also was resampled randomly 500 times. This was done by generating a single random vector from a multivariate normal distribution whose mean vector was β , as calculated from the original data, and whose covariance matrix was the matrix of sample variances (squared standard errors) and covariances of the original estimates of β . The predicted response to selection was calculated for each resampling of genetic and selection parameters, and its standard error obtained as the standard deviation of the replicates.

RESULTS

The First Selection Episode

Selection in 1976–1977.—Prior to the period of selection, 634 birds measured in 1975 and 1976 were alive in mid-1976. Of these, 15% ($N = 98$) survived to 1978 when selection had ceased (Boag and Grant 1981, 1984a). Survivors were larger than nonsurvivors in all six dimensions (t -tests, $P < 0.001$ in all cases). Most birds were of unknown sex. Small samples of male ($N = 24$) and female ($N = 14$) sur-

TABLE 1. Directional natural selection on *Geospiza fortis* in 1976–1977 and 1984–1986. Coefficients in boldface are significantly different ($P < 0.05$) from zero.

| | Standardized selection coefficients | | |
|---------------|-------------------------------------|---------------|-------|
| | Differential | Gradient | |
| | s | β | SE |
| 1976–1977 | | | |
| Weight | +0.74 | +0.477 | 0.146 |
| Wing length | +0.72 | +0.436 | 0.126 |
| Tarsus length | +0.43 | +0.005 | 0.110 |
| Bill length | +0.54 | –0.144 | 0.174 |
| Bill depth | +0.63 | +0.528 | 0.214 |
| Bill width | +0.53 | –0.450 | 0.197 |
| Sample size | 634 | 632 | |
| Survival | | 15% | |
| 1984–1986 | | | |
| Weight | –0.11 | –0.040 | 0.101 |
| Wing length | –0.08 | –0.015 | 0.084 |
| Tarsus length | –0.09 | –0.047 | 0.076 |
| Bill length | –0.03 | +0.245 | 0.095 |
| Bill depth | –0.16 | –0.135 | 0.136 |
| Bill width | –0.17 | –0.152 | 0.125 |
| Sample size | 556 | 549 | |
| Survival | | 32% | |

vivors exhibited the same trend displayed by the total sample and by the sample of birds of unknown sex, although not all individual traits showed significant mean increases.

Standardized selection differentials and gradients are given in table 1. The net effects of selection, as measured by selection differentials, were strongest on weight and wing length, which are two independent measures of structural size (Grant et al. 1985). Trait means shifted from approximately one-half to three-quarters of a standard deviation. Gradient analysis shows that net effects on trait means were brought about by direct selection for large weight, wing length, and bill depth. These direct and independent effects of selection more than counterbalanced selection in the opposite direction on one other trait, bill width. Thus, not only was large overall size selectively favored, but bill shape, that is depth in relation to width, also changed under selection.

Results of the gradient analysis are robust to deletion of single variables, with one exception (see also Price and Boag 1987). Deletion of either bill depth or bill width eliminates the significant entry in the gradient of the other. This comes about from their high intercorrelation ($r = 0.857$); all other pairs of variables are less strongly but positively correlated (0.4–0.8). The problem of multicollinearity (high character correlations) is well recognized in selection analyses (Lande and Arnold 1983; Mitchell-Olds and Shaw 1987) and is often dealt with by eliminating one of a pair of correlated variables (e.g., Price and Boag 1987, Grant and Grant 1993). In the present analysis, both bill traits are retained because, even though they are highly correlated, each has *significant* and *opposite* independent effects on fitness (survival).

These results confirm and extend previously reported analyses of the same selection event (Boag and Grant 1981, 1984a; Price and Boag 1987; Price et al. 1984b;). Trivial differences in the magnitude (but not signs) of the selection

TABLE 2. Evolutionary response of *Geospiza fortis* to selection in 1976–1977. All differences between generations (Δ) are shown in boldface, indicating that they are significant by *t*-tests; for tarsus length, $P = 0.0048$, for all others $P < 0.0001$.

| | Before selection 1976 | | Next generation 1978 | | Δ |
|--------------------|--------------------------|------|-------------------------|------|--------------|
| | Mean | SE | Mean | SE | |
| Weight (g) | 16.06 | 0.06 | 17.13 | 0.13 | +1.07 |
| Wing length (mm) | 67.88 | 0.10 | 68.87 | 0.20 | +0.99 |
| Tarsus length (mm) | 19.08 | 0.03 | 19.29 | 0.07 | +0.21 |
| Bill length (mm) | 10.63 | 0.03 | 10.95 | 0.06 | +0.32 |
| Bill depth (mm) | 9.21 | 0.03 | 9.70 | 0.06 | +0.49 |
| Bill width (mm) | 8.58 | 0.02 | 8.83 | 0.05 | +0.25 |
| Sample size | 634 | | 135 | | |

coefficients arise from reclassification of eight individuals (Grant 1993), new measurements of a few more, and inclusion of wing length and tarsus length in the analyses. Addition of these two variables has two effects; wing length becomes an extra significant entry into the selection gradient, without affecting the others, and the total r^2 rises from 0.105 to 0.122.

Evolutionary Response.—Heritabilities of all six morphological traits are high and significantly different from zero. They vary from 0.48 ± 0.13 for tarsus length to 0.97 ± 11 for weight. Similarly, genetic correlations between traits are all large and positive, varying from 0.67 ± 0.15 for tarsus length and weight to 0.94 ± 0.10 for bill depth and width. Therefore, evolutionary responses to strong directional selection are expected to be pronounced in the cohort of offspring born in 1978, and are in fact observed (table 2). All differences between the sample before selection and the sample of members of the next generation are significant at $P < 0.005$.

These results are not affected by a possible bias arising from sexual selection on the parental generation. Sexual selection on male traits could have occurred in this basically monogamous species because males outnumbered females among the survivors by at least 2:1 (Boag and Grant 1981, 1984a); because they were larger than females on average, males survived better when large size was selectively favored (Gibbs and Grant 1987b). An unbalanced sex ratio gives females an opportunity to choose mates on the basis of traits previously subjected to natural selection (Price 1984; Grant and Grant 1989a,b). If they chose mates in this way, the breeders would be a nonrandom sample of the survivors. This did not happen: the males that bred did not differ in any trait from those that did not breed (*t*-tests, $P > 0.1$); hence, there is no evidence of sexual selection on any of the traits.

High heritabilities imply that offspring born in 1978 to the survivors will be close in size to their parents, although a little smaller on average. This expectation was realized; offspring had smaller means than their parents, though none of the differences were significant (all $P > 0.1$).

We repeated the analysis of the evolutionary response by using members of the 1981 cohort; the 1979 and 1980 cohorts of measured birds were too small for separate analysis. Although sexual selection occurred in 1981 (Price 1984), all of the previously selected adults alive that year bred and produced a total of 16 measured offspring, which we compared

TABLE 3. Comparison of predicted and observed evolutionary responses of *Geospiza fortis* to selection in 1976–1977. There are no significant differences between the two by *t*-tests. The observed means were calculated from the restricted sample of 36 offspring hatched in 1978 whose parents were included in the selection analysis. They differ from means of the unrestricted sample used in table 2.

| | Predicted | | Observed | | Δ in SD units | |
|---------------|-----------|------|----------|------|----------------------|----------|
| | Mean | SE | Mean | SE | Predicted | Observed |
| Weight | 17.39 | 0.22 | 17.53 | 0.25 | +0.82 | +0.91 |
| Wing length | 69.98 | 0.39 | 69.65 | 0.35 | +0.85 | +0.72 |
| Tarsus length | 19.45 | 0.09 | 19.32 | 0.14 | +0.48 | +0.32 |
| Bill length | 11.14 | 0.10 | 11.06 | 0.11 | +0.68 | +0.58 |
| Bill depth | 9.83 | 0.12 | 9.94 | 0.09 | +0.76 | +0.90 |
| Bill width | 8.96 | 0.08 | 8.97 | 0.08 | +0.65 | +0.67 |

with the sample of 634 birds before selection. Tarsus length was not included because it was measured differently in 1981 (T. D. Price pers. comm. 1981). Four of the traits differed significantly ($P < 0.001$), confirming the original result. In addition, mean wing length was 1 mm larger in the 1981 offspring than in the preselected group, but the difference is not significant ($P = 0.1051$). The 16 offspring did not differ in any trait from 18 offspring of the same parents ($N = 7$ families) born in the years 1979 and 1980 combined ($P > 0.25$ in all cases). Thus, the observed evolutionary response in 1978 was a repeatable consequence of selection and not simply a product of exceptionally favorable conditions for growth in 1978.

Observed and Predicted Responses Compared.—Table 3 shows a comparison of offspring means with those predicted from Lande's (1979) equation for multivariate evolution. The sample used to calculate the observed evolutionary response comprised the 36 (measured) offspring hatched in 1978 whose parents were banded before selection; other measured offspring had at least one parent that did not enter the selection analysis and might, for this reason, not be characteristic of the sample used in the selection analysis.

There is good agreement between observed and predicted values, no significant differences between any pair, and no trend to under- or overpredict the means. As with the total sample, all observed means are significantly different from the sample before selection; therefore, all traits show a significant positive response.

The Second Selection Episode

Selection in 1984–1986.—Before selection occurred, 556 birds measured before mid-1984 were alive. Ignoring a brief and largely unsuccessful attempt by a few birds to breed in 1986, the next breeding season occurred in 1987. Of the starting group of 556, 180 (32%) survived to 1987. The net effect of selection up to this time was a decrease in the mean size of traits, especially bill dimensions (table 1).

In the gradient analysis of all six traits, only one, bill length, was revealed as a target of selection; moreover, an increase in this trait was selectively favored. Net effects of selection, size decreases, were produced by the cumulative small effects in that direction on all other traits. This becomes clearer when traits are deleted singly from the gradient anal-

TABLE 4. Evolutionary response of *Geospiza fortis* to selection in 1984–1986. Differences between generations (Δ) shown in boldface are significant ($P < 0.001$) by t -tests; for the others $P > 0.05$.

| | Before selection 1984 | | Next generation 1987 | | Δ |
|--------------------|--------------------------|------|-------------------------|------|--------------|
| | Mean | SE | Mean | SE | |
| Weight (g) | 17.04 | 0.07 | 16.45 | 0.07 | -0.59 |
| Wing length (mm) | 68.17 | 0.10 | 67.08 | 0.10 | -1.09 |
| Tarsus length (mm) | 19.10 | 0.03 | 19.02 | 0.03 | -0.08 |
| Bill length (mm) | 10.92 | 0.03 | 10.99 | 0.03 | +0.07 |
| Bill depth (mm) | 9.61 | 0.03 | 9.36 | 0.03 | -0.25 |
| Bill width (mm) | 8.86 | 0.02 | 8.74 | 0.02 | -0.12 |
| Sample size | 556 | | 434 | | |

ysis. Results are robust to deletion with the exception of bill depth and width, as in the 1976–1977 selection analysis. These are the two dimensions most strongly correlated. Deletion of bill depth results in the coefficient for bill width becoming significant ($\beta = -0.251 \pm 0.101$, $P = 0.013$), and deletion of bill width results in the coefficient for bill depth becoming significant ($\beta = -0.236 \pm 0.094$, $P = 0.012$). In each case, the significance of the β value for bill length remains unchanged, and none of the other β values change to a large degree or become significant. Notice that the sum of the two coefficients in table 1 is almost the same as the individual ones after deletion.

These results are almost the same as those reported earlier (Gibbs and Grant 1987a) from analyses of a slightly smaller data set for a shorter time (1984–1985). Several small errors in the original data set have been corrected (see also Grant and Grant 1993).

Evolutionary Response.—Observed evolutionary responses were calculated from the trait means of 378 measured offspring hatched in 1987 whose measured parents were included in the “before-selection” group. Excluded were another 58 measured offspring with only one parent in the preselection group. All responses were significant ($P < 0.001$) except tarsus length and bill length ($P > 0.05$, table 4). They all agree in direction with the signs of the β coefficients, that is, offspring were smaller on average than in the preselected group, except that bill length gives an indication of a shift in the opposite direction. Inclusion of the 58 offspring mentioned above does not alter any of the results.

There is almost no scope for the potential bias, considered in the 1976–1977 selection episode, of sexual selection on the parental generation. The sex ratio of survivors was 1:1, and all but three individuals bred. Of the survivors, 99% had been measured, and only 10 (6%) were measured after the middle of 1984 and hence were excluded from the preselected group.

Estimates of the evolutionary response did not depend on conditions unique to 1987. The next full breeding season was 1991, another wet year of extensive breeding like 1987 and with a similar breeding density (Grant and Grant 1992). There is no a priori reason to expect growth conditions to have been different in the two years. Seven pairs produced measured offspring in both 1987 (25) and 1991 (11). Although the statistical power to detect small differences between these

TABLE 5. Comparison of predicted and observed evolutionary responses of *Geospiza fortis* to selection in 1984–1986. Figures in boldface show where observed offspring generation means differ significantly from predicted means by t -tests that allow for unequal variances (Sokal and Rohlf 1981); for weight $P < 0.05$, wing length $P < 0.001$, and bill depth $P < 0.01$. The observed means were calculated from the restricted sample of 378 offspring hatched in 1987 whose parents were included in the selection analysis. They differ from means of the unrestricted sample used in table 4.

| | Predicted | | Observed | | Δ in SD units | |
|---------------|-----------|------|--------------|------|----------------------|----------|
| | Mean | SE | Mean | SE | Predicted | Observed |
| Weight | 16.82 | 0.13 | 16.48 | 0.08 | -0.13 | -0.33 |
| Wing length | 67.93 | 0.17 | 67.21 | 0.11 | -0.11 | -0.41 |
| Tarsus length | 19.02 | 0.04 | 19.02 | 0.04 | -0.08 | -0.10 |
| Bill length | 10.86 | 0.05 | 10.96 | 0.03 | -0.10 | +0.05 |
| Bill depth | 9.51 | 0.06 | 9.32 | 0.03 | -0.14 | -0.39 |
| Bill width | 8.77 | 0.04 | 8.70 | 0.03 | -0.15 | -0.27 |

two groups is not large, they do not differ significantly in any trait, in either a comparison of all individuals or a comparison of family means ($P > 0.1$ in all cases).

Observed and Predicted Responses Compared.—Unlike the outcome of the 1976–1977 selection analysis, observed and predicted responses to the second episode of selection differ significantly in three traits (table 5). With the exception of bill length, which displayed no evolutionary response (table 4), observed means decreased more than predicted (table 5). The observed decreases (mean SD shift = 0.24) are twice as large as those predicted (mean = 0.12 SD).

Comparison of the 1978 and 1987 Evolutionary Responses

None of the observed and predicted responses in 1978 differ significantly, whereas three of the six differ significantly in 1987. The superiority of the first predictions is partly but not entirely due to the broader confidence limits on the estimates of offspring trait means in 1978 resulting from smaller sample sizes than those available in 1987. In fact, the predictions are closer to observed responses in 1978 than in 1987 (fig. 1). This can be quantified by comparing the distance between the observed vector of standardized evolutionary change and the predicted vector in the two years (tables 3, 5). The distance, calculated as the square root of the sum of the squared differences for the six traits, is 0.28 in 1978 and 0.48 in 1987.

The multivariate equation does a better job of prediction in 1978 than in 1987 in two other respects. First, the vector correlation between observations and predictions is stronger in 1978 (0.86) than in 1987 (0.42). Second, whereas all observed means in 1978 are correctly predicted to be larger than those of the preselected group, one of the 1987 means (bill length) ostensibly shifted in the opposite direction to that predicted. Furthermore, selection for small size in the parental generation (1984–1986), coupled with heritabilities of less than 1.0, should have led to the 1987 offspring being larger than their parents, but in four traits they were smaller. There was no corresponding discrepancy in the 1978 prediction.

We now examine four hypotheses to account for the main

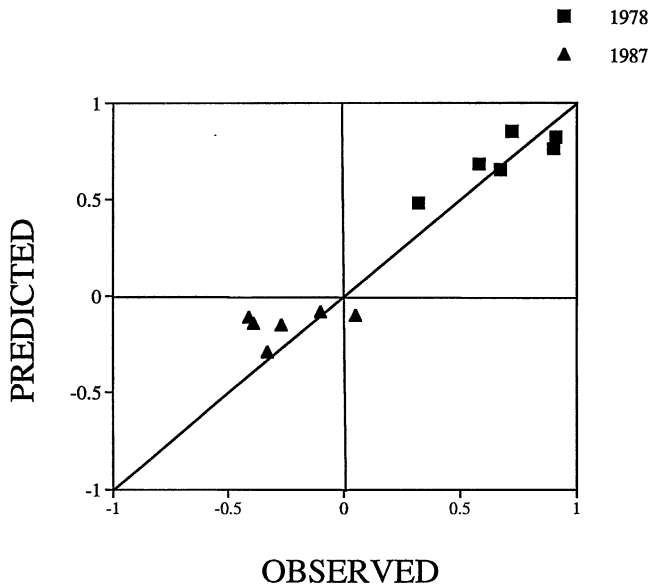


FIG. 1. Comparison of predicted and observed evolutionary responses of *Geospiza fortis* to selection in 1976–1977 and 1984–1986. The responses were manifested by the next generations born in 1978 and 1987, respectively. Data are taken from tables 3 and 5.

differences between predictions and observations in 1987, three of which were formally significant.

Hypotheses to Explain Departures of Observations from Predictions

Incorrect Estimation of Genetic Parameters.—The large bill-length mean observed in the 1987 offspring is anomalous. Given the relative magnitudes of the β coefficients and the strengths of the positive genetic correlations among traits, the upward selection on bill length should have been more than outweighed by the cumulative effects of downward selection on all of the other traits. A possible explanation for the anomaly is that incorrect genetic values have been used. The anomaly would diminish or disappear if the heritability of bill length and genetic correlations with other traits were close to zero. This is not observed. Most of the breeders in 1987 were born in 1983. Bill-length heritability calculated from this cohort of offspring born in 1983 is 0.58 ± 0.13 ($N = 62$ families), which is similar to the estimate of 0.65 ± 0.15 for the earlier cohorts (Boag 1983). Heritabilities for other bill traits are higher (>0.60), and all genetic correlations are positive and large (>0.60). These are much too high to account for either the anomalously high mean bill length or the low mean values of other traits in the offspring born in 1987.

Natural Selection on the Offspring.—Given that parameters of the multivariate predictive equation have been correctly estimated, the source of any discrepancy with observations probably lies in two factors affecting the offspring generation. First, before being measured, this cohort may have been subjected to directional natural selection. Second, members of this cohort may have experienced conditions during their growth to adult size that were different from those experienced by their parents as they grew to maturity. Although

selective mortality need not be independent of nutritional effects on growth and adult size, we first consider and test two hypotheses that treat them separately before examining a possible interaction. Finally, offspring and parents may have been subjected to different environmental influences after they had stopped growing.

Selection for small size early in life, which has been inferred for other cohorts (Price and Grant 1984), could have occurred before most of the 1987 cohort were measured. Selection for large bill length could have occurred at the same time. In other words, the offspring may have been subjected to the same selective pressures experienced by the parental generation in previous years. Assuming selection mainly affects between-brood variation, the hypothesis predicts that the measured 1987 cohort members should have smaller than average parents because those inheriting large size from large parents should have suffered relatively heavy mortality.

The hypothesis was tested by comparing the breeders that produced measured offspring ($N = 147$) with those that did not ($N = 10$). The analysis was restricted to those families in which both parents were measured and were members of the 1984 preselection group. None of the single-trait comparisons were significant ($P > 0.1$) by t -tests, nor were there any suggestive trends: the producers had smaller means than the nonproducers for only wing and tarsus length. The hypothesis is not supported by results of this test.

The test is limited in sensitivity by the small sample size of parents that contributed no offspring to the measured sample. As an alternative, we regressed the number of measured offspring ($N = 0$ –12) produced by parents on the midparent values of the six measured traits. Since offspring number is a measure of fitness, the analysis was performed as a selection analysis, with relative fitness regressed on standardized ln-transformed variables. The resulting regression was not significant ($F = 1.016$; $df = 6, 150$; $P = 0.4173$), and none of the individual β coefficients was significant. Thus, the hypothesis is not supported. However, the two largest coefficients also predominated in the 1984–1986 selection gradient, and had the same sign. For bill length $\beta = 0.103 \pm 0.078$ ($t = 1.32$, $P = 0.190$), and for bill width $\beta = -0.115 \pm 0.096$ ($t = 1.20$, $P = 0.2331$). Therefore, differential production of surviving offspring among parents of different bill sizes (and shapes) could have affected the measurement of the evolutionary response, but only to a small extent.

These analyses were repeated with the 1978 cohort, with similar (negative) results. Ten pairs produced one to seven measured offspring. The overall regression was not significant ($F = 0.703$; $df = 6, 3$; $P = 0.6745$), and none of the individual β coefficients approached significance ($P > 0.25$).

Nutritional Effects on Growth.—Poor growth in 1987, as apparently occurred in 1976 (Boag 1983, 1987), could have led to below-average size of birds born that year. This would not explain the relatively high bill length mean, but positive selection for this trait may have occurred, despite generally poor conditions for growth, just as it did in the parental generation. Even though food supply was plentiful in 1987, an El Niño year, initial breeding density was high at ≈ 100 pairs (Grant and Grant 1992). Conditions for growth may have been inferior to those experienced by their parents. Most parents were born (hatched) in 1982–1983, another El Niño

TABLE 6. Mean measurements of offspring hatched in the first half, January through April (early, $N = 184$), and in the second half, May through August (late, $N = 194$), of the breeding season in 1987. Boldface shows where the (only) early offspring mean differs significantly ($P < 0.05$) from the predicted mean (table 5) by a t -test.

| | Early offspring | | Late offspring | | t | P |
|---------------|-----------------|------|----------------|------|-------|-------|
| | Mean | SE | Mean | SE | | |
| Weight | 16.67 | 0.10 | 16.30 | 0.11 | 2.479 | 0.014 |
| Wing length | 67.46 | 0.15 | 66.97 | 0.15 | 2.338 | 0.020 |
| Tarsus length | 18.97 | 0.05 | 19.06 | 0.05 | 1.166 | 0.244 |
| Bill length | 10.94 | 0.05 | 10.97 | 0.05 | 0.592 | 0.554 |
| Bill depth | 9.38 | 0.04 | 9.25 | 0.05 | 2.002 | 0.046 |
| Bill width | 8.73 | 0.04 | 8.67 | 0.04 | 1.377 | 0.169 |

year, when initial breeding density was only 70 pairs. The hypothesis makes two interrelated predictions. Offspring born in 1987 should be smaller, on average, than their parents, and the difference should be greater in the second half of the breeding season, at a time of higher density, than in the first half. Note that selection for small size in the parental generation, combined with heritabilities of less than 1.0, should lead to offspring being larger, not smaller, than their parents in the absence of the hypothesized nutritional effects.

The first prediction was tested by comparing offspring ($N = 378$) with their parents ($N = 147$). For three of the six traits differences were significant, and in each case they were in agreement with expectation; offspring were significantly smaller than their parents in weight ($t_{521} = 2.53$, $P = 0.0118$), wing length ($t_{521} = 3.46$, $P = 0.0006$) and bill depth ($t_{523} = 2.92$, $P = 0.0036$). Offspring mean bill width ($\bar{x} = 8.70 \pm 0.02$ mm) was smaller than the parental mean ($\bar{x} = 8.76 \pm 0.05$), but not to a significant extent ($t_{523} = 1.23$, $P = 0.2187$). Offspring means for the other two traits, bill length and tarsus length, were almost identical to, but actually larger than, the corresponding parental means as is to be expected from their heritabilities of less than one.

The second prediction was tested by comparing the sizes of offspring born early (January–April, $N = 184$ measured) and late (May–August, $N = 194$) in the 1987 breeding season (table 6). Results were almost the same as those from the first test and supported the hypothesis. Late-born birds were significantly smaller than early born birds in weight, wing length, and bill depth. A difference in bill width in the same direction is not significant. Differences in the opposite direction for tarsus and bill length are small and not significant.

Thus, the growth hypothesis is supported by tests of two predictions. The explanation for tarsus length not decreasing may be that almost all growth is completed in the nest, unlike all the other traits (Boag 1984), and that factors preventing the attainment of potential adult size such as insufficient food and competitive interactions operate after fledging. The same explanation cannot be applied to bill length because this trait takes the longest time to reach adult size.

The higher survival of small birds could also be produced by selection for small relative size within broods, independent of absolute size (i.e., between broods). We have no evidence for this, and results of our only relevant comparison run counter to it. Recaptured offspring had higher average

weights at banding (day 8) than their nest mates not recaptured (Wilcoxon Matched-Pairs Signed-Rank test, $z = 2.67$, $N = 152$, $P = 0.0076$ two-tailed). Nonetheless, selection for small size within broods might occur after fledging. This would imply that the smallest member of each brood is in better condition on average than the rest. In this way, selective mortality and differential growth may interact and be, strictly, inseparable.

To conclude, a role for nutritional factors affecting growth and causing discrepancies between observed and predicted evolutionary responses in 1987 seems certain. A role for selection for small offspring size remains possible, and for large bill length is probable.

To reduce the hypothesized density-mediated environmental effects on growth and adult size of the 1987 cohort, we use only the early-born birds instead of the total sample for comparison with predicted means. This substitution (table 6) results in a closer correspondence between observed and predicted trait means and the disappearance of two of the three significant differences. Only wing length remains significantly smaller than predicted (t 's = 2.07, $P < 0.05$). The observed decreases (mean SD shift = 0.19) are closer to those predicted (mean = 0.12 SD) than are those for the total sample of offspring (mean = 0.24 SD).

With regard to the 1978 cohort, although they were larger than their parents in all traits, as predicted, mean values for four of the traits fell slightly short of predicted means. A nutritional effect on final adult size may have been responsible, but is unlikely because conditions for growth that year were favorable: breeding density was low (Boag 1983), and nestling food supply (arthropods) was plentiful (Boag and Grant 1984b). Moreover, offspring born early ($N = 59$) and late ($N = 66$) in the 1978 season did not differ significantly from each other in any of the six traits (t -tests, $P > 0.05$ in each case).

Direct Effects of the Environment.—Offspring hatched in 1987 experienced different conditions from those of their parents after they had attained adult size. Two drought years followed 1987 (Grant and Grant 1992). None of their parents experienced such a long period with limited opportunities for molting their plumage before 1987. Therefore, the wing-length comparison is likely to be affected by a bias arising from abrasion of feathers and low molting frequency during the years when most of the offspring were measured. This could account for the significant difference between observed and predicted wing lengths.

DISCUSSION

The desirability of making testable evolutionary predictions from the measurement of selection in the wild contrasts with the rarity of the opportunity to do so. Predictions require estimation of the strengths and directions of selection on traits, and quantitative genetic parameters of those traits. Predictions can be tested against observations when offspring of identified parents in the selected group have been measured. Such offspring measurements are often difficult if not impossible to obtain. We have taken advantage of the convenience of a small and resident population of birds to predict

evolution in some morphological traits, to measure evolution in those traits, and to compare the two (fig. 1).

We found good agreement between observations and predictions. Some discrepancies in the second episode stimulated us to investigate possible biases. We identified a biasing environmental influence on adult traits in one cohort: a combination of food supply and density affected the adult size attained by growth. These effects may have been enhanced by weight loss and feather abrasion during an unusually long drought (1988–1990) before some of the birds were measured (see also Gosler 1987), and by selection on the offspring generation prior to their measurement. Similar problems are likely to be frequently encountered in any study conducted in fluctuating environments, which are just those in which selection is expected to occur repeatedly.

Univariate and Multivariate Analyses of Selection

There are general lessons to be learned from these attempts to predict evolution, over and above the value they have in contributing to an understanding of contemporary selection pressures on the particular species. The value of the multivariate approach to the evolution of quantitative traits, and the need to keep selection separate from evolution, are well illustrated by both analyses of selection. In neither episode did selection affect only one trait. Correlated characters were selected as a set, and the genetic correlations determined the evolutionary consequences. In most cases, traits subjected to selection underwent evolutionary change. However, bill length was not a target of selection in 1976–1977; nevertheless, it evolved as a correlated response to selection on other traits. In contrast, bill length was selected in the 1984–1986 episode but it did not evolve, partly because the effect was nullified by selection in the opposite direction on positively correlated traits. Evolutionary change in this dimension, or lack of it, could not be understood without a knowledge of the correlation structure.

Boag (1983) found a high correlation ($r_s = 0.89$) between the elements of the phenotypic and genetic correlation matrices. All elements are large. When \mathbf{G} and \mathbf{P} , the genetic and phenotypic covariance matrices, are almost the same, the predicted responses to selection on the various traits approach the selection differentials, because $\Delta\mathbf{z} = \mathbf{GP}^{-1}\mathbf{S}$ (Lande 1979). We may therefore ask, How well does the standardized version of the simple predictive equation $R = h^2s$ perform? The answer is, surprisingly well.

The vector correlation between predicted and observed standardized responses is actually much higher in 1987 when this simple equation is used (0.70) than when the multivariate equation is used (0.42); for 1978, the vectors for the simple (0.92) and multivariate (0.86) equations are nearly identical. However, there is a loss in the accuracy of some individual predictions. Bill depth predicted by the univariate equation in 1978 is significantly lower ($P < 0.05$) than observed. In fact, the univariate equations perform less well than the multivariate equations in terms of the closeness of predictions to observations (as opposed to the pattern of correspondence between the two). The vector of distances between predictions and observations in 1978 is lower for the multivariate

equations (0.28) than it is for the univariate equations (0.43). The vectors are more nearly equal in 1987; 0.48 for the univariate equation and 0.44 for the multivariate equation. For most organisms and most characters, with lower heritable variation and hence higher environmental variation (see Boag and van Noordwijk 1987, Roff 1990), genetic and phenotypic matrices will correspond less well and the simple predictive equation will be less reliable. This is a subject for empirical investigation.

Statistical and Biological Assumptions

When natural selection is relatively weak, as will usually be the case, the power of the predictive method of evolutionary study will be limited by imprecise and incomplete parameter estimation. In particular, the success of multivariate analysis is dependent upon statistical assumptions being met. Statistical assumptions of selection analyses have been discussed extensively (Lande and Arnold 1983; Endler 1986; Manly 1985; Price and Boag 1987; Mitchell-Olds and Shaw 1987, 1990; Crespi and Bookstein 1989; Crespi 1990; Wade and Kalisz 1990; Price and Langen 1992; Rauscher 1992; van Tienderen and de Jong 1994). A potentially serious problem is the failure to include all relevant variables. This can result in misidentification of the true targets of selection and a concomitant failure to predict its evolutionary consequences correctly. For example, evolutionary responses of bill length and depth in 1987 were not accurately predicted, possibly because another, correlated trait not included in the analysis also was subjected to natural selection. We do not know of any trait that is positively correlated with one of the bill dimensions and negatively correlated with the other. This explanation would require that some genetic and phenotypic correlations differ greatly, which is not known for any pair of variables in our studies. Problems arising from unidentified targets of natural selection are not confined to field studies and may occur even in well-controlled laboratory studies of artificial selection (e.g., Beniwal et al. 1992).

The problem of missing variables can never be dealt with completely. Many studies of natural selection in the wild are open-ended and are likely to suffer acutely from the missing-variables problem. Crespi (1990) has urged a maximally inclusive analysis to minimize the problem. Our study is different in being focused on a small set of beak and body size variables that were chosen for measurement (Grant et al. 1976; Abbott et al. 1977) because they had been identified previously as important in the morphological differentiation and speciation of the group (Swarth 1934; Lack 1945, 1947; Bowman 1961). Moreover, a combination of known food-supply changes and a biomechanical argument for the significance of bill shape adequately accounts for the observed directional selection and predictable evolutionary response (we have not discussed the causes of selection in this paper; see Grant 1986; Grant and Grant 1989a).

The main biological assumption of the analyses is correct identification of parents. Misidentification gives rise to error when the true parents are not included in the selection analysis and when they differ systematically in the selected traits from those that are included.

Some studies of related (emberizine) finches in continental

regions have found extrapair fertilizations and hence cryptic paternity to be moderately frequent (e.g., Petter et al. 1990; Westneat et al. 1990), as determined by DNA fingerprinting; others have found them to be rare (Hartley et al. 1993). In our analyses, we assume we have correctly identified paternity and maternity by observation. Although no molecular data are available to test the assumption, it is supported by two facts. First, observations on extrapair copulations are extremely rare in the week before and during the egg-laying period (Price 1984, unpubl. obs.), when the female is presumed to be receptive (Birkhead and Møller 1992), and egg laying in the nests of other females has never been observed. Second, regressions of offspring measurements on the measurements of the apparent father are no different from those on measurements of the apparent mother. Both sets of regressions are very high (Boag 1983). Therefore, if the true biological parents have been misidentified, only a small number of families could have been affected by this type of error.

Long-Term Evolution

These findings on short-term evolutionary responses to selection have important implications for our understanding of long-term evolution. There is a well-established tradition of making simple extrapolations from the former to the latter (e.g., Futuyma 1986 and others). As a recent example, Williams (1992) calculated the average strength of selection each generation necessary to produce long-term evolutionary change, assuming a constant supply of genetic variation, and shows that large changes can be brought about over a long time by a very small amount of selection per generation. He argued that recurrent mutation is likely to replenish genetic variance lost through selection (e.g., see Lande 1975; Ayala 1982; Charlesworth 1987). He also pointed to long-term selection experiments like those of Yoo (1980), which show undiminished responses to persistent directional selection, though small in magnitude at each generation, and concluded that the "logical difficulty with additive genetic variation and responsiveness to selection is that there is too much, not too little" (Williams 1992, p. 64). Earlier, Price et al. (1984a) had employed the methods of this paper, explicitly dealing with character correlations, to estimate the minimal selective forces involved in transitions from one Darwin's finch species to another. The results of the 1976–1977 selection episode were extrapolated to speciation. Schluter (1984) extended the method to reconstruction of the finch phylogeny.

The results of our study caution that the realized selective forces involved in evolutionary transitions are likely to be greater than the extrapolated theoretical minima. It is probably more realistic to consider selection acting episodically and changing in sign, within and between generations (Price and Grant 1984; Schluter and Smith 1986b), yet resulting in a net trend over the long term. For example, Hairston and Walton (1986) and Hairston and Dillon (1990) demonstrated a pattern of fluctuating selection on the timing of diapause by freshwater copepods (*Diaptomus sanguineus*) caused by predatory fish. The timing trait is heritable (in the lab), changes annually, and displays an apparent evolutionary response to selection. There are many examples in the fossil record of fluctuating patterns of morphological trait variation,

and, although it is difficult to rule out ecophenetic change, they are consistent with a regime of oscillating selection on a broad time scale (e.g., Kellogg 1975; Malmgren et al. 1983; Reymont 1983; MacLeod 1991).

Our results show an additional reason why simple extrapolations may be faulty. The evolutionary response to an episode of directional selection may be less than predicted for environmental reasons. Environmental (nutritional) effects may prevent potential size of the offspring generation from being realized. Deviations from the long-term average of environmental conditions are likely to affect the offspring generation more than the parental generation, because the offspring generation is a single cohort whereas the parents comprise several cohorts. For example, in our study parents of the 1987 offspring were born (hatched) in the years 1975, 1978, 1979, 1981, 1983, and 1984, years of very different rainfall, breeding density, and food supply (Grant and Grant 1992). Interannual variation in environmental influences on adult size are averaged in the parental generation. For semelparous organisms this factor will have more pronounced effects.

A third and similar reason has been identified and discussed by Price et al. (1988), van Noordwijk et al. (1988), and Alatalo et al. (1990) (see also Rauscher 1992; van Tienderen and de Jong 1994). This is the possibility that selection acts on the environmental deviations of size-related characters subject to food limitation during growth in the parental generation and is therefore ineffective in causing evolution.

These three reasons, in addition to the familiar one of constraints on change arising from genetic correlations, answer the question of why evolutionary responses over a long period of time might be much less than expected by simply extrapolating from measured genetic variation and selection in the short term. As such, they are relevant to other discussions of long-term evolution, such as apparent stasis punctuated by directional change (Gould and Eldredge 1993), and the design of tests that attempt to distinguish between drift and directional selection as causes of evolution on the basis of simple rates of divergence (Cheetham et al. 1993; Spicer 1993).

There are important differences between the two selection regimes we have discussed. Even though the endpoint may be the same, the cumulative forces of selection will be much greater if they oscillate in direction than if the direction is constant and if they act partly on the environmental component of variance. The time course will be longer. However, the opposite would be true if genetic correlations, when constraining evolution, are modified more easily by oscillating than by unidirectional selection. The long-term dynamics of genetic variances and covariances under different regimes of selection are not well enough understood for this possibility to be assessed.

To some extent, the distinction we have made between oscillating and unidirectional selection and response depends on temporal scale. Unidirectional microevolutionary change over the short term becomes just part of oscillating change over a longer period of time. For example, climates naturally fluctuate seasonally and annually about a long-term average. The evolutionary changes caused by oscillating directional selection arising from changes in rainfall in our study can be

considered a model for long-term evolution caused by climatic fluctuations; strong evolutionary change brought about by one climatic extreme was reversed by evolutionary change brought about by a contrasting climatic extreme. The distinction is not just a matter of scale, however, and sustained unidirectional change can occur in many ways, such as persistent responses to the challenges of new predators, parasites, pathogens, or competitors (e.g., Seeley 1986). Climatic change may play a role in causing such changes in community composition, and thereby have indirect as well as direct effects upon the ecological and evolutionary fates of species.

ACKNOWLEDGMENTS

We thank the Dirección General de Desarrollo Forestal, Quito, the Galápagos National Park Service, and the Charles Darwin Research Station for permission to conduct the field research, Natural Sciences and Engineering Research Council (Canada) and National Science Foundation for financial support, D. Schluter for performing the bootstrap analyses, S. Pacala for discussion, and R. Gomulkiewicz, T. D. Price, D. Schluter, and an anonymous reviewer for comments on the manuscript.

LITERATURE CITED

- Abbott, I. L. K., L. K. Abbott, and P. R. Grant. 1977. Comparative ecology of Galápagos ground finches (*Geospiza* Gould): evaluation of the importance of floristic diversity and interspecific competition. *Ecological Monographs* 47:151–184.
- Alatalo, R. V., L. Gustafsson, and A. Lundberg. 1990. Phenotypic selection on heritable size traits: environmental variance and genetic response. *American Naturalist* 135:464–471.
- Ayala, F. J. 1982. The genetic structure of species. Pp. 60–82 in R. Milkman, ed. *Perspectives on evolution*. Sinauer, Sunderland, Mass.
- Beniwal, B. K., I. M. Hastings, R. Thompson, and W. G. Hill. 1992. Estimation of changes in genetic parameters in selected lines of mice using REML with an animal model. II. Body weight, body composition and litter size. *Heredity* 69: 361–371.
- Birkhead, T. R., and A. P. Møller. 1992. *Sperm competition in birds. Evolutionary causes and consequences*. Academic Press, New York.
- Boag, P. T. 1983. The heritability of external morphology in Darwin's ground finches (*Geospiza*) on Isla Daphne Major, Galápagos. *Evolution* 37:877–894.
- . 1984. Growth and allometry of external morphology in Darwin's finches (*Geospiza*) on Isla Daphne Major, Galápagos. *Journal of Zoology*, London 204:413–441.
- . 1987. Effects of nestling diet on growth and adult size of Zebra finches (*Poephila guttata*). *Auk* 104:155–166.
- Boag, P. T., and P. R. Grant. 1978. Heritability of external morphology in Darwin's finches. *Nature* 274:793–794.
- . 1981. Intense natural selection in a population of Darwin's finches (*Geospizinae*) in the Galápagos. *Science* 214:82–85.
- . 1984a. The classical case of character release: Darwin's finches (*Geospiza*) on Isla Daphne Major, Galápagos. *Biological Journal of the Linnean Society* 22:243–287.
- . 1984b. Darwin's finches (*Geospiza*) on Isla Daphne Major, Galápagos: breeding and feeding ecology in a climatically variable environment. *Ecological Monographs* 54:463–489.
- Boag, P. T., and A. J. van Noordwijk. 1987. Quantitative genetics. Pp. 45–78 in F. Cooke and P. A. Buckley, eds. *Avian genetics. A population and ecological approach*. Academic Press, New York.
- Bowman, R. I. 1961. Morphological differentiation and adaptation in the Galápagos finches. *University of California Publications in Zoology* 58:1–302.
- Bulmer, M. G. 1980. *The mathematical theory of quantitative genetics*. Clarendon Press, Oxford.
- Charlesworth, B. 1987. The heritability of fitness. Pp. 21–40 in J. W. Bradbury and M. B. Andersson, eds. *Sexual selection: testing the alternatives*. Wiley, Chichester, U.K.
- Cheetham, A. H., J. B. C. Jackson, and L.-A. Hayek. 1993. Quantitative genetics of bryozoan phenotypic evolution. I. Rate tests for random change versus selection in differentiation of living species. *Evolution* 47:1526–1538.
- Cheverud, J. M. 1984. Quantitative genetics and developmental constraints on evolution by selection. *Journal of Theoretical Biology* 110:155–171.
- Crespi, B. J. 1990. Measuring the effect of natural selection on phenotypic interacting systems. *American Naturalist* 135:32–47.
- Crespi, B. J., and F. L. Bookstein. 1989. A path-analytic model for the measurement of selection on morphology. *Evolution* 43:18–28.
- Endler, J. A. 1986. *Natural selection in the wild*. Princeton University Press, Princeton, N.J.
- Falconer, D. 1989. *Introduction to quantitative genetics*, 3d ed. Longman, Harlow, Essex, U.K.
- Flux, J. E. C., and M. M. Flux. 1982. Artificial selection and gene flow in wild starlings, *Sturnus vulgaris*. *Naturwissenschaften* 69: 96–97.
- Futuyma, D. J. 1986. *Evolutionary biology*, 2d ed. Sinauer, Sunderland, Mass.
- Gibbs, H. L., and P. R. Grant. 1987a. Oscillating selection on Darwin's finches. *Nature* 327:511–513.
- . 1987b. Adult survival in Darwin's ground finch (*Geospiza*) populations in a variable environment. *Journal of Animal Ecology* 56:797–813.
- Gosler, A. G. 1987. Pattern and process in the bill morphology of the great tit *Parus major*. *Ibis* 129:451–476.
- Gould, S. J., and N. Eldredge. 1993. Punctuated equilibrium comes of age. *Nature* 366:223–227.
- Grant, B. R., and P. R. Grant. 1989a. Natural selection in a population of Darwin's finches. *American Naturalist* 133:377–393.
- . 1989b. Evolutionary dynamics of a natural population. The large cactus finch of the Galápagos. University of Chicago Press, Chicago.
- . 1993. Evolution of Darwin's finches caused by a rare climatic event. *Proceedings of the Royal Society of London B* 251: 111–117.
- Grant, P. R. 1986. *Ecology and evolution of Darwin's finches*. Princeton University Press, Princeton, N.J.
- . 1993. Hybridization of Darwin's finches on Isla Daphne Major, Galápagos. *Philosophical Transactions of the Royal Society of London, B* 340:127–139.
- Grant, P. R., and B. R. Grant. 1992. Demography and the genetically effective sizes of two populations of Darwin's finches. *Ecology* 63:766–784.
- . 1994. Phenotypic and genetic effects of hybridization in Darwin's finches. *Evolution* 48:297–316.
- . 1995. The founding of a new population of Darwin's finches. *Evolution* 49:229–240.
- Grant, P. R., B. R. Grant, J. N. M. Smith, I. J. Abbott, and L. K. Abbott. 1976. Darwin's finches: populations variation and natural selection. *Proceedings of the National Academy of Sciences, USA* 73:257–261.
- Grant, P. R., I. Abbott, D. Schluter, R. L. Curry, and L. K. Abbott. 1985. Variation in the size and shape of Darwin's finches. *Biological Journal of the Linnean Society* 25:1–39.
- Gustafsson, L. 1989. Collared Flycatcher. Pp. 75–88 in I. Newton, ed. *Lifetime reproduction in birds*. Academic Press, New York.
- Hairston, N. G., Jr., and T. A. Dillon. 1990. Fluctuating selection and response in a population of freshwater copepods. *Evolution* 44:1796–1805.
- Hairston, N. G., Jr., and W. E. Walton. 1986. Rapid evolution of a life history trait. *Proceedings of the National Academy of Sciences, USA* 83:4831–4833.
- Haldane, J. B. S. 1937. *Adventures of a biologist*. Harper and Bros., New York.
- Hartley, I. R., M. Shepherd, T. Robson, and T. Burke. 1993. Re-

- productive success of polygynous male corn buntings (*Miliaria calandra*) as confirmed by DNA fingerprinting. *Behavioral Ecology* 4:310–317.
- Kellogg, D. 1975. The role of phyletic change in the evolution of *Pseudocubus vema* (Radiolaria). *Paleobiology* 1:359–370.
- Kingsolver, J. G., and D. C. Wiernasz. 1991. Development, function, and the quantitative genetics of wing melanin pattern in *Pieris* butterflies. *Evolution* 45:1480–1492.
- Kirkpatrick, M., and D. Lofsvold. 1992. Measuring selection and constraint in the evolution of growth. *Evolution* 46:954–971.
- Lack, D. 1945. The Galápagos finches (*Geospizinae*): a study in variation. *Occasional Papers of the California Academy of Science* 21:1–159.
- . 1947. Darwin's finches. Cambridge University Press, Cambridge.
- Lande, R. 1975. The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genetical Research* 26:221–235.
- . 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution* 33:402–416.
- Lande, R., and S. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* 37:1210–1226.
- MacLeod, N. 1991. Punctuated anagenesis and the importance of stratigraphy to paleobiology. *Paleobiology* 17:167–188.
- Malmgren, B. A., W. A. Berggren, and G. P. Lohmann. 1983. Evidence for punctuated gradualism in the late Neogene *Globorotalia timida* lineage of planktonic foraminifera. *Paleobiology* 9:377–384.
- Manly, B. F. J. 1985. The statistics of natural selection. Chapman and Hall, London.
- Mitchell-Olds, T., and R. G. Shaw. 1987. Regression analysis of natural selection: statistical inference and biological interpretation. *Evolution* 41:1149–1161.
- . 1990. Comments on the causes of natural selection. *Evolution* 44:2158.
- Petter, S. C., D. B. Miles, and M. M. White. 1990. Genetic evidence of mixed reproductive strategy in a monogamous bird. *Condor* 92: 702–708.
- Price, T. D. 1984. Sexual selection on body size, territory and plumage variables in a population of Darwin's finches. *Evolution* 38:327–341.
- Price, T. D., and P. T. Boag. 1987. Selection in natural populations of birds. Pp. 257–287 in F. Cooke and P. A. Buckley, eds. *Avian genetics. A population and ecological approach*. Academic Press, New York.
- Price, T. D., and P. R. Grant. 1984. Life history traits and natural selection for small body size in a population of Darwin's finches. *Evolution* 38:483–494.
- Price, T. D., and T. Langen. 1992. Evolution of correlated characters. *Trends in Ecology and Evolution* 7:307–310.
- Price, T. D., P. R. Grant, and P. T. Boag. 1984a. Genetic changes in the morphological differentiation of Darwin's Ground finches. Pp. 49–66 in K. Wöhrmann and V. Loeschcke, eds. *Population biology and evolution*. Springer, Berlin.
- Price, T. D., P. R. Grant, H. L. Gibbs, and P. T. Boag. 1984b. Recurrent patterns of natural selection in a population of Darwin's finches. *Nature* 309:787–789.
- Price, T. D., M. Kirkpatrick, and S. J. Arnold. 1988. Directional selection and the evolution of breeding date in birds. *Science* 240:798–799.
- Rauscher, M. D. 1992. The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. *Evolution* 46:616–626.
- Reyment, R. A. 1983. Phenotypic evolution in microfossils. Pp. 209–254 in M. K. Hecht, B. Wallace, and G. T. Prance, eds. *Evolutionary biology*, vol. 16. Plenum, New York.
- Roff, D. A. 1990. Understanding the evolution of insect life cycles: the role of genetic analysis. Pp. 5–27 in F. Gilbert, ed. *Insect life cycles. Genetics, evolution and co-ordination*. Springer, Berlin.
- Schluter, D. 1984. Morphological and phylogenetic relations among the Darwin's finches. *Evolution* 38:921–930.
- Schluter, D., and J. N. M. Smith. 1986a. Genetic and phenotypic correlations in a natural population of song sparrows. *Biological Journal of the Linnean Society* 29:23–36.
- . 1986b. Natural selection on beak and body size in the song sparrow. *Evolution* 40:221–231.
- Seeley, R. H. 1986. Intense natural selection caused a rapid morphological transition in a living marine snail. *Proceedings of the National Academy of Sciences, USA* 83:6897–6901.
- Sokal, R. R., and F. J. Rohlf. 1981. *Biometry*. W. H. Freeman, San Francisco.
- Spicer, G. S. 1993. Morphological evolution of the *Drosophila virilis* species group as assessed by rate tests for natural selection on quantitative characters. *Evolution* 47:1240–1254.
- Swarth, H. S. 1934. The bird fauna of the Galápagos Islands in relation to species formation. *Biological Reviews* 9:213–234.
- van Noordwijk, A. J., and J. H. van Balen. 1988. The great tit, *Parus major*. Pp. 119–135 in T. H. Clutton-Brock, ed. *Reproductive success*. University of Chicago Press, Chicago.
- van Noordwijk, A. J., J. H. van Balen, and W. Scharloo. 1988. Heritability of body size in a natural population of the great tit and its relation to age and environmental conditions during growth. *Genetical Research* 51:149–162.
- van Tienderen, P. H., and G. de Jong. 1994. A general model of the relation between phenotypic selection and genetic response. *Journal of Evolutionary Biology* 7:1–12.
- Wade, M. J., and S. Kalisz. 1990. The causes of natural selection. *Evolution* 44:1947–1955.
- Westneat, D. F., P. W. Sherman, and M. L. Morton. 1990. The ecology and evolution of extra-pair copulations in birds. *Current Ornithology* 7:331–369.
- Williams, G. C. 1992. *Natural selection: domains, levels, and challenges*. Princeton University Press, Princeton, N.J.
- Yoo, B. H. 1980. Long-term selection for a quantitative character in large replicate populations of *Drosophila melanogaster*. I. Response to selection. *Genetical Research* 35:1–17.

Corresponding Editor: D. Schluter